DOCUMENT RESUME

UD 031 947

ED 412 322

AUTHOR Mac Iver, Douglas J.; Plank, Stephen B.; Balfanz, Robert

TITLE Working Together To Become Proficient Readers. Early Impact of the Talent Development Middle School's Student Team Literature Program. Report No. 15.

INSTITUTION Center for Research on the Education of Students Placed At Risk, Baltimore, MD.; Howard Univ., Washington, DC.; Johns Hopkins Univ., Baltimore, MD.

SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.; Carnegie Corp. of New York, NY.

PUB DATE 1997-08-00

NOTE 38p.

CONTRACT R-117-D40005

PUB TYPE Reports - Evaluative (142)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS Curriculum; Disadvantaged Youth; *Educational Change; *Helping Relationship; Intermediate Grades; Junior High Schools; *Literature; Mentors; *Middle Schools; Peer Relationship; Reading Achievement; Reading Comprehension; *Reading Programs; School Restructuring; Teaching Methods; Urban Schools

IDENTIFIERS Hierarchical Linear Modeling

ABSTRACT

The Talent Development Model of Middle School Reform includes a "Student Team Literature" (STL) program that relies on: (1) curricular materials designed to assist students to study great literature; (2) recommended instructional practices, peer assistance processes, and assessments; and (3) staff development, mentoring, and advising to support the curricular and instructional reforms. In February of 1995, Central East Middle School in Philadelphia (Pennsylvania) became the United States' first school to adopt the Talent Development Model. Data on students' prior reading achievement, achievement after the first year of implementation, and the frequency of peer assistance were collected in 21 STL classes and in 25 comparison classes in a closely matched control school. Hierarchical linear modeling analyses that controlled for prior reading achievement reveal that students in STL classes display significantly better reading comprehension after the first year of implementation (effect size = .51). Although the impact of participating in STL on students' reading comprehension is sizable for students across the entire prior achievement spectrum, students with the strongest prior reading skills especially benefited. In addition, peer assistance was more frequent and productive in STL classes than in comparison classes. An appendix contains the student evaluation questionnaire items. (Contains 4 tables, 7 figures, and 47 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 412 322

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to: UD

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view. CS

CRESPAR

WORKING TOGETHER TO BECOME PROFICIENT READERS

Early Impact of the Talent Development Middle School's *Student Team Literature Program*

Douglas J. Mac Iver • Stephen B. Plank • Robert Balfanz

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Report No. 15 / August 1997

CENTER FOR RESEARCH ON THE EDUCATION OF STUDENTS PLACED AT RISK

Johns Hopkins University & Howard University

BEST COPY AVAILABLE

WORKING TOGETHER TO BECOME PROFICIENT READERS

Early Impact of the Talent Development Middle School's *Student Team Literature Program*

Douglas J. Mac Iver, Stephen B. Plank, and Robert Balfanz
Johns Hopkins University

Report No. 15

August 1997

Published by the Center for Research on the Education of Students Placed At Risk (CRESPAR), supported as a national research and development center by funds from the Office of Educational Research and Improvement, U.S. Department of Education (R-117-D40005). Additional funding for this work was received from the Carnegie Corporation of New York. The opinions expressed in this publication do not necessarily reflect the position or policy of OERI or the Carnegie Corporation, and no official endorsement should be inferred.

The Center

Every child has the capacity to succeed in school and in life. Yet far too many children, especially those from poor and minority families, are placed at risk by school practices that are based on a sorting paradigm in which some students receive high-expectations instruction while the rest are relegated to lower quality education and lower quality futures. The sorting perspective must be replaced by a "talent development" model that asserts that all children are capable of succeeding in a rich and demanding curriculum with appropriate assistance and support.

The mission of the Center for Research on the Education of Students Placed At Risk (CRESPAR) is to conduct the research, development, evaluation, and dissemination needed to transform schooling for students placed at risk. The work of the Center is guided by three central themes — ensuring the success of all students at key development points, building on students' personal and cultural assets, and scaling up effective programs — and conducted through seven research and development programs and a program of institutional activities.

CRESPAR is organized as a partnership of Johns Hopkins University and Howard University, in collaboration with researchers at the University of California at Santa Barbara, University of California at Los Angeles, University of Chicago, Manpower Demonstration Research Corporation, WestEd Regional Laboratory, University of Memphis, and University of Houston-Clear Lake.

CRESPAR is supported by the National Institute on the Education of At-Risk Students (At-Risk Institute), one of five institutes created by the Educational Research, Development, Dissemination and Improvement Act of 1994 and located within the Office of Educational Research and Improvement (OERI) at the U.S. Department of Education. The At-Risk Institute supports a range of research and development activities designed to improve the education of students at risk of educational failure because of limited English proficiency, poverty, race, geographic location, or economic disadvantage.

Abstract

The Talent Development Middle School's *Student Team Literature* (STL) program includes: (1) curricular materials designed to assist students study great literature; (2) recommended instructional practices, peer assistance processes, and assessments; and (3) staff development, mentoring, and advising to support the curricular and instructional reforms. Data on students' prior reading achievement, achievement after the first year of implementation, and on the frequency of peer assistance were collected in 21 STL classes and in 25 comparison classes in a closely matched control school. HLM analyses that control for prior reading achievement reveal that students in STL classes display significantly better reading comprehension after the first year of implementation (effect size=.51). Although the impact of participating in STL on students' reading comprehension is sizable for students across the entire prior achievement spectrum, students with the strongest prior reading skills especially benefit. In addition, peer assistance is more frequent and more productive in STL classes than in comparison classes.

Acknowledgments

The authors wish to thank Leslie Jones, Maria Garriott, Salimah Perkins, Alan Ruby, and Bob Petza for their vital contributions. We are grateful to David Hornbeck, the superintendent of the School District of Philadelphia, and the teachers, administrators, and students at Central East Middle School in Philadelphia for making this work possible. The authors also wish to thank Gary Natriello, Sharon Robinson, and John Hollifield for their comments on an earlier version of this report.

Introduction

> The best predictors of student achievement that are within a school's control are the content actually taught, the instructional strategies used, and the standards for achievement evident in testing and grading. — Porter, 1993, p. 2

In March 1997, officers from four private foundations that support middle school reform initiatives in 19 states and more than 600 schools reported that the middle school reform movement has reached a critical juncture (Lipsitz, Mizell, Jackson, & Austin, 1997). As a result of foundation-supported reform efforts, many middle schools have changed their climates and structures to become "warmer, happier, and more peaceful places for both students and adults.... [However, most of these schools] have not moved off this plateau and taken the critical next step to develop students who perform well academically, with the intellectual wherewithal to improve their life conditions (p. 535)." The collective experience of Lipsitz and her colleagues with four major middle-level school reform initiatives has convinced them that the middle school reform movement may ultimately fail unless more schools move further "along the developmental continuum from changing climate and structure toward changing curriculum and instruction" (p. 540) so as to create academically excellent centers of teaching and learning that strive for every student's mastery of a demanding curriculum.

Fostering Academic Achievement in the Middle Grades

The existing literature on student achievement indicates that in order to create an environment in which every middle school student can achieve a high level of intellectual proficiency four interrelated aspects of the social organization of learning need to be improved: the curriculum, instructional and peer assistance strategies, assessment, and professional development.

The Curriculum

Middle school students need to be engaged by a common core curriculum which is coherent, focused, and challenging. Simply put, students usually do not learn what they are not taught or given a structured opportunity to acquire (Knapp, 1995). In addition, there is growing evidence that student engagement is higher when academic tasks are more complex

(Stodolsky, 1988; Bruer, 1994). Student engagement in turn has been shown to be a key element in the rate and level of learning (Means, Chelemer, & Knapp, 1991). Unfortunately, international comparisons, national surveys, and case studies all indicate that the middle grades in the United States currently fall short on one and often all of these features (Balfanz, 1997; Corbett & Wilson, 1997; Knapp, 1995; Stevenson & Stigler, 1992; U.S. Department of Education, 1996). Tracked middle schools typically deny students in the lower tracks the opportunity to acquire and master higher order competencies (Oakes, Gamoran, & Page, 1992). The existing middle school curriculum is often fragmented and repetitive. As a result, students are not afforded the opportunity to acquire new knowledge in an organized and systematic fashion. In the recent Third International Mathematics and Science Study (TIMSS), for example, U.S. students ranked 23rd out of 27 in the rate of growth of mathematical knowledge between seventh and eighth grade (Beaton, et al., 1996). Researchers attribute much of this low ranking to a curriculum which is "a mile wide and an inch thick" (Schmidt, McKnight, & Raizen, 1996). Both the earlier work of Stevenson and Stigler and the more recent TIMSS research argue that high levels of academic achievement are attained in nations where classroom instruction is supported by curricula which introduce advanced skills and concepts to all middle grade students in a thoughtful, focused, and active manner.

Instructional Strategies

By itself, however, a coherent, focused, and challenging core curriculum is not enough to dramatically increase the number of middle school students achieving a high level of academic excellence. Detracking, inclusion, and the changing demographics of the student population in the United States have led to middle school classrooms which are populated with students who are from more diverse backgrounds and who have a wider range of prior academic success and preparation than in the past (Natriello, McDill, & Pallas, 1990). Heterogeneous secondary school classrooms, for instance, can contain students at five or more instructional levels (Mevarech & Kramarski, 1997). This shift in the distribution of student characteristics at the classroom level from a more narrow range of instructional levels and prior experiences to a more varied one has increased the complexity of middle school instruction. Fuchs, Fuchs, Mathes, and Simmons (1997) argue that in the absence of new instructional strategies which enable teachers to effectively structure learning in this more complex environment, teachers will simplify the instructional challenge by focusing on the students they consider "teachable" and eliminating the "difficult to teach students from consciousness." In other words, unless the organization of instruction in middle schools is

expanded to include a repertoire of instructional strategies that are effective in heterogeneous classes, the elimination of tracking and the institution of a common core curriculum may still not improve the learning of a sizable number of students.

Fortunately, a number of instructional strategies have shown promise in delivering high-level instruction to heterogeneous classes. These include, but are not limited to, reciprocal teaching (Palinscar & Klenk, 1991), classwide peer tutoring (Fuchs et al., 1997), the East Asian whole class method (Stevenson & Lee, 1995), and several variations of cooperative learning (King, 1994; Mevarech & Kramarski, 1997; Slavin, 1995; Webb & Farivar, 1994). The common threads which run through these approaches are the use of peer-assisted learning, explicit mechanisms for providing students with essential background knowledge, an emphasis on developing metacognitive strategies, and materials or strategies which engage students in an active way with questions which provoke higher order thinking. What remains to be shown is that these strategies or other effective methods of teaching in heterogeneous classes can be implemented and sustained across the core subjects and throughout the middle grades.

Assessment

The existing literature suggests that increasing the number of middle school students who acquire advanced academic skills will also require several reforms in assessment. First, few points have been as clearly made and abundantly disseminated in the past decade than the need to align classroom instruction and assessment (e.g., American Federation of Teachers, 1992; Glaser & Silver, 1994). Yet in practice, because of the multiplicity of authorities to whom schools, teachers, and students are accountable, this remains a difficult task. Second, at the classroom level, students need to be provided with continual corrective feedback on their work (Locke & Latham, 1990, chapter 8). Seatwork, however, which has traditionally occupied much of students' time in the middle grades, commonly goes unchecked and uncorrected (Stevenson & Stigler, 1992). Third, assessments at the classroom level need to facilitate or at least support a classroom environment in which every student is encouraged and motivated to put forth solid effort (Mac Iver & Reuman, 1993; Mac Iver, Reuman, & Main, 1995). Fourth, assessments should feature performance tasks that are "essential, integrative, rich, engaging, active, and feasible." (Forgione, 1990, p. 3).

Professional Development

Sustained, schoolwide, and curriculum-specific professional development is needed to realize the improvements sought in curriculum, instructional strategies, and assessment in the middle grades. But there is only limited research, so far, on what constitutes effective professional development (McLaughlin & Oberman, 1996), although this research has made several key points. First, Cohen, Wilson, and Hill (1997) in a survey of elementary mathematics teachers in California found that generic staff development (sessions on using manipulatives or conducting standards-based mathematics lessons, for example) had no effect on teacher self-reports of classroom practice or district achievement scores. Curriculum-specific sessions, on the other hand, which provided teachers with content knowledge (operating with fractions, for example) and which prepared them to teach actual replacement units incorporating standards-based lessons, had a positive effect on both teacher self-reports of classroom practice and classroom achievement. The authors state: "This is a terrifically important result, if it stands up, for it suggests that when teachers have significant opportunities to learn the content that students will learn in ways that seem to enable them to learn more about teaching the material — and when assessments are linked to the students' and teachers' curriculum — teachers' opportunities to learn pay off for their students' learning" (p. 61). Another finding of note is the importance of continual technical assistance and follow-up (Fuchs et al., 1997). Finally, there is growing evidence of the need to involve the entire teaching staff in a focused and collaborative effort to improve teaching and learning in a school (Nelson & Hammerman, 1996; Elmore, Peterson, & McCarthey, 1996). Professional development efforts which target a handful of teachers or particular subject in a particular grade may lead to "pockets of excellence" but they do not create successful learning communities (Knapp, 1995).

The challenge then for the middle grades is not simply to improve curriculum, instruction, assessment, and professional development but to do it throughout the school, across grades, and across subjects. This is a tall order. It is made even taller by the fact that it needs to occur not only in schools where the conditions for reform may be favorable, but also in all other middle schools, including those often found in our nation's inner cites, which face the extreme conditions of large class sizes, high mobility, and high concentrations of poverty (Knapp, 1995). This evaluation study reports on the first-year accomplishments of an effort to foster reading comprehension in an inner-city middle school where teachers face extreme conditions (class sizes of 33 and over 85 percent low-income families) by changing the

curriculum, instructional strategies, assessments, and professional development in Reading, English, and Language Arts (RELA). This RELA-focused effort is Phase I of a three-year comprehensive plan to reform the curriculum, instruction, school organization, and support services of the entire school, addressing all major subject areas.

The Talent Development Model of Middle School Reform

In February 1995, teachers and administrators at Central East Middle School in Philadelphia voted to become the nation's first school to adopt the Talent Development Model of middle school reform. This model stresses "access to opportunity" reforms and is guided by the belief that all students can learn challenging academic material if the proper resources, support, and school organization are provided (Mac Iver & Plank, 1998; Madhere & Mac Iver, 1996). All students are exposed to a common core curriculum that is coherent, standards-based, and aimed at active student learning of higher-order competencies. Other components of the model include heterogeneous grouping of students for all core academic classes; expanded opportunities for extra academic help and enrichment during the regular school day; expanded opportunities for academic recognition; a communal organization of the school that includes semi-departmentalized staffing with interdisciplinary teams and small learning communities that endure for two or three years; the establishment of strategic school-family-community partnerships that emphasize communicating, learning at home, involving parents in school decisions, and collaborating with the community; an emphasis on promoting cultural literacy; and weekly career exploration and educational decision-making lessons that assist students in setting goals, planning for the future, and systematically exploring educational and career options. Central East began a three-year process of phasing in these components in September 1995. During the first year of this process, Central East Middle School focused on helping its students become more proficient readers of outstanding literature by implementing the Talent Development Middle School's *Student Team Literature* program in Reading, English and Language Arts (RELA). In September 1995, Central East also instituted heterogeneous grouping in all major subject areas, a communal organization of the school, and elective "extra dose" classes during the regular school day attended by those needing extra academic help.

The Talent Development Model's *Student Team Literature* Program

The Student Team Literature program (Jones, 1997) is an adaptation and elaboration of the Student Team Reading program (Stevens, 1989; Stevens & Durkin, 1992). It includes (1) curricular materials, (2) recommended instructional practices, peer assistance processes, and assessments, and (3) staff development, mentoring, and advising to support the curricular and instructional reforms. We believe that Student Team Literature may prove to be even more effective than Student Team Reading in raising students' reading comprehension for several reasons. Student Team Literature (1) focuses on great books rather than short selections from an anthology; (2) eliminates literal comprehension questions from partner discussion guides and literature tests and replaces them with higher-level questions; (3) greatly reduces the use of out-of-context reading comprehension worksheets; (4) makes teacher's guides more comprehensive (e.g., in providing background information and suggesting specific listening comprehension activities to enhance students' understanding of the literary devices used by the author); and (5) introduces more variety in the materials, instructional practices, and processes to prevent monotony.

Curricular materials. The curricular materials produced by the Student Team Literature Program are designed to assist students and teachers as they study great literature together. They consist of *teacher guides,* and students' *partner discussion guides*, *literature tests*, and *word mastery tests* for over 60 award-winning novels and plays. The partner discussion guides present students with challenging, high-level questions about the literature they are reading.[1] The questions require them to think analytically and comparatively, to draw conclusions, and to analyze the plot, literary devices, and the writer's style and technique. For example, students are asked to explain the significance of the main characters and of the setting in relation to the story, to draw conclusions about any special relationships between characters, to make predictions regarding what will happen next based on solid textual evidence that they have already read, and to develop interesting, unusual, or thought-provoking ideas related to the literature's theme. In addition, each section of a partner discussion guide has a word mastery list so that students may work with their teammates to master the vocabulary used in a novel or play; "lit review" activities to assist teammates in preparing each team member to do well on the short literature test given after each section of

---

[1] Talent Development Program curriculum writers at Johns Hopkins University write the partner discussion guides with Bloom's taxonomy of educational objectives in mind (Bloom, 1984). The guides are intended to stimulate thinking consistent with each of the six cognitive objectives — knowledge, comprehension, application, analysis, synthesis, and evaluation — with a special focus on the last four of these.

a novel or play; writing prompts to guide students in writing a brief composition related to one of the topics or themes of the literature they are studying, and suggested extension activities. The partner discussion guides are designed to be used in conjunction with Student Team Literature's instructional practices and cooperative learning processes (discussed below) and help to ensure a schoolwide focus in RELA on higher-order competencies.

Each teacher guide includes a summary of the novel or play and information about its author. It also includes suggested pre-reading activities to help students acquire or recall relevant background information that will enhance their understanding and appreciation of the literature, discussion questions for use in leading whole-class discussions, suggested listening comprehension (read-aloud) activities to help students identify and interpret the literary devices used by the author, and suggested answers to every question in students' partner discussion guides.

Instructional practices, processes, and assessment. In Student Team Literature, as in all Student Team Learning methods (Slavin, 1995), students work in cooperative teams. Each team is made up of four to five students of different races, genders, and levels of achievement, and each reflects the composition of the class as a whole on these attributes. The teams receive awards and recognition based on the degree to which they can ensure the high academic performance of each member on literature tests, word mastery tests, and essays. The team awards give students a reason to support the achievement efforts of their peers because the team can succeed only if each team member learns.

One of the distinctive features of a class using Student Team Literature is the structured and systematic way students work together (with their partners and with the rest of their cooperative learning team) to assess and support each other's learning as they read, comprehend, discuss, and analyze literature. Specific activities within Student Team Literature include: (1) partner reading, in which students read part of a novel silently and then read the same part aloud with a partner; (2) partner discussion, in which students discuss answers to questions in the partner discussion guide and then individually write their answers to these questions; (3) word mastery, in which students practice saying new vocabulary words aloud with their partners and also write vivid "meaningful sentences" to show that they understand meanings; (4) pretest activities, in which students assess each teammate's comprehension of the novel selection by asking each other "story retell" questions from the partner discussion guide, give each other a "words out loud" pretest to assess whether each teammate can read the vocabulary words smoothly and easily, and assess each other's mastery of word meanings by reviewing teammates' meaningful sentences; and (5) story-related essay writing (followed by peer review, then by revision, editing, and publication).

In Student Team Literature, classroom assessment is seamlessly aligned with instruction. Students are assessed each week by their teammates (the immediate systematic feedback from peers that was described in the previous paragraph) and by their teacher (formal individual assessments upon which both individual grades and team scores are based). The weekly individual assessments include: (1) a Literature Test (students write short answers to five questions such as, "Why are the Nisei and Issei reluctant to sign the questionnaire forswearing allegiance to the emperor of Japan?"); (2) a Word Meaning Test (students write meaningful sentences for high frequency, highly useful words from the literature selection such as "warily"); (3) an oral Words Out Loud Test given individually by the teacher to determine students' ability to read the new vocabulary words from the literature selection smoothly and easily; and (4) the final draft of a student's story-related essay (sample writing prompt: "Write a three-page essay comparing *The Invisible Thread* to another book you have read. For example, compare and contrast the "prison" experience of Yoshiko to that of Anne Frank; compare the racial prejudice that affected Yoshiko with that which affected Anne Frank or Harriet Tubman. Compare and contrast the way Yoshiko faces insurmountable odds with the way another protagonist — such as Harriet Tubman, Anne Frank, or Johnny Tremain — deals with conflict and difficulties.").

As with most effective cooperative learning methods of instruction, Student Team Literature also involves daily direct instruction by the teacher. This includes vocabulary instruction (e.g., familiarizing students with words found in their literature selection while modeling decoding strategies, defining words, modeling their use in sentences, and leading the class in composing meaningful sentences for some of the high frequency, highly useful words), introducing the literature selection and building background, and leading whole-class discussion of essential points from the literature selection.

Student Team Literature also includes a regularly scheduled Listening Comprehension time, during which students are read to by a model reader — their teacher — for about twenty minutes (Jones, 1996; Stevens & Shaw, 1991). During this time, students are able to hear a fluent reader read and think aloud. Before commencing the reading, the teacher introduces a literary element or device. She or he reads a selection in which the element or device is prominently featured. While reading, the teacher pauses occasionally to ask questions of students or to think aloud for them. Finally, the teacher leads a brief concluding discussion, again highlighting the featured literary device. The activity is designed to strengthen students' comprehension skills, teach new literary elements and devices, and improve students' listening skills.

Staff development. The staff development, mentoring, and advising which supported the curricular and instructional reforms began in June of 1995 and continued throughout the 1995-96 school year. On summer days, and on scheduled staff development weekdays during the school year, one or more Student Team Literature developers would describe and discuss an aspect of the RELA curricular and instructional reforms. A central part of these staff development sessions was usually a modeling time in which the Student Team Literature developer would play the role of the teacher while Central East staff members would play the roles of students assigned to cooperative teams.

Other mentoring and advising activities included trainers observing a classroom and then holding a coaching session with the teacher to discuss successes and failures, to review key parts of the RELA reforms, and to suggest strategies that might be used in the future. Finally, developers met periodically with small groups of teachers to review the materials from the various staff development days. Most teachers attending these review sessions had missed the initial presentations or desired further exposure to them.

In March of the school year during which this evaluation study was conducted, RELA teachers at Central East Middle School also received staff development in Student Team Writing (Stevens, Hammonds, & Durkin, 1991). In Student Team Writing, students work cooperatively when planning, revising, and editing their writing. They learn to give feedback to one another and to use feedback from one another while preparing their writing for publication (in classroom compilations, in "author's chair" presentations to their classmates, in writing competitions, on the World Wide Web, and so on).

Previous Research Evaluating the Effectiveness of Student Team Reading

As mentioned, Student Team Literature is an adaptation and elaboration of Student Team Reading. Previous research has been conducted on the effects of using Student Team Reading during the middle grades to teach RELA using a literature anthology. Stevens and Durkin (1992) conducted two evaluations of Student Team Reading in urban middle schools. The first evaluation examined the use of Student Team Reading in 20 experimental sixth grade classes in three schools matched with 39 classes in three control schools. Experimental students achieved significantly higher on a standardized measure of reading comprehension (effect size=.11). Also, the reading comprehension achievement of academically handicapped students, analyzed separately, was significantly higher in the experimental group.

Stevens and Durkin's second evaluation examined the use of Student Team Reading and Student Team Writing in sixth, seventh, and eighth grade classes in two urban middle schools matched with three control schools. In this study, students being taught using these cooperative learning methods had significantly higher measures of reading vocabulary, reading comprehension (effect size=.25), and language expression.

Stevens and Durkin's studies provide convincing evidence of the effectiveness of Student Team Reading in raising students' reading comprehension. It was our hope that our adaptation and elaboration of this technique would prove even more effective in raising students' reading comprehension. We also hoped that the Student Team Literature program would increase substantially the frequency of systematic, focused, and effective peer academic assistance and discussion in RELA classrooms. Stevens and Durkin (1992) did not directly measure Student Team Reading's impact on how frequently and how effectively students helped other students learn. Given that virtually every middle school teacher is repeatedly urged these days to give his or her students frequent opportunities to learn cooperatively, it is important to investigate Student Team Literature's success in increasing the frequency and effectiveness of students' efforts to assist their peers.

Method

Participants

Our analyses involve all RELA classes that serve sixth, seventh, and eighth graders at two schools in Philadelphia (21 classes at Central East Middle School and 25 classes at our control site). Ungraded students in self-contained special education classes are not included. After Central East had agreed to become the nation's first Talent Development Middle School, the research office of the School District of Philadelphia identified the other school as an appropriate comparison site, because it was very similar to Central East in terms of student population and characteristics of the teaching staff. Neither school is a magnet school. Each school serves fifth, sixth, seventh, and eighth grades.[2] Each enrolls approximately 1,000 students. Over 85 percent of the students from each school come from low-income families. Also, many of the students have learned English as their second language. At Central East, 53 percent of the students spoke a language other than English before they started going to school. At the control school, the comparable number is 48 percent.

---

[2] The analyses in this report include only sixth, seventh, and eighth grade students because we do not have prior test scores (from Spring, 1995) for fifth graders.

Regarding racial and ethnic composition, Central East's student body is about 45 percent Hispanic, with most of these students being of Puerto Rican descent. Another 24 percent of the students are African American. About 13 percent of the students are of Asian descent, primarily Cambodian and Vietnamese. Another 8 percent are white. The remaining 10 percent identify themselves as belonging to other racial or ethnic groups. These other groups include Arab Americans and biracial students.

The control school's student body is similarly diverse although, in comparison with Central East, it has a higher proportion of Hispanic and white students, a somewhat lower proportion of African American students, and very few Asian American students. Specifically, 57 percent of the students are Hispanic; 16 percent are African American; 15 percent are white; fewer than 1 percent are Asian American; and the remainder describe themselves as biracial or other.

Data Collection and Measures

Measure of peer assistance in RELA. As part of a larger questionnaire that was administered to all students in both schools in February 1996, students responded to a four-item scale measuring the frequency of peer assistance and discussion in their RELA class. By computing the mean student response within each classroom to these items, we were able to create a scale that reliably measured the frequency with which peer assistance activities had occurred in RELA class during the first five months of the school year. The items and alpha coefficient for this scale are presented in the appendix.

In the majority of cases, the surveys were administered in the RELA classroom, with both the teacher and a Johns Hopkins University researcher present. In other cases, due to scheduling constraints, surveys were administered in non-RELA classes, but with the instructions that students should answer the questions with reference to RELA. A total of 1,301 sixth, seventh, and eighth graders were surveyed, 664 from the control school and 628 at Central East Middle School.

Measures of reading comprehension and of prior reading achievement. Reading comprehension was operationalized as a student's Reading Comprehension Scale Score from the Stanford 9's multiple choice battery. This test was administered in April 1996, near the end of the first year of Student Team Literature implementation at Central East Middle School. Our measure of prior achievement is the reading total scale score from the CTBS achievement test, administered in May 1995. Although the use of two different tests means

that we cannot precisely measure gains in achievement, the CTBS score serves as a reliable covariate measuring achievement prior to the implementation of Student Team Literature.

Results

Our analysis plan was (1) to use dummy variable multiple regression to estimate the difference between experimental and control classrooms in our classroom-level measure of peer assistance, (2) to use hierarchical linear models (HLM) to estimate the difference between experimental and control classrooms in the students' reading comprehension while controlling for prior achievement and current grade level, and (3) to use HLM to estimate differences in the effectiveness of peer assistance in increasing reading comprehension in experimental and control classrooms.

School as a Predictor of Between-Classroom Differences in Peer Assistance

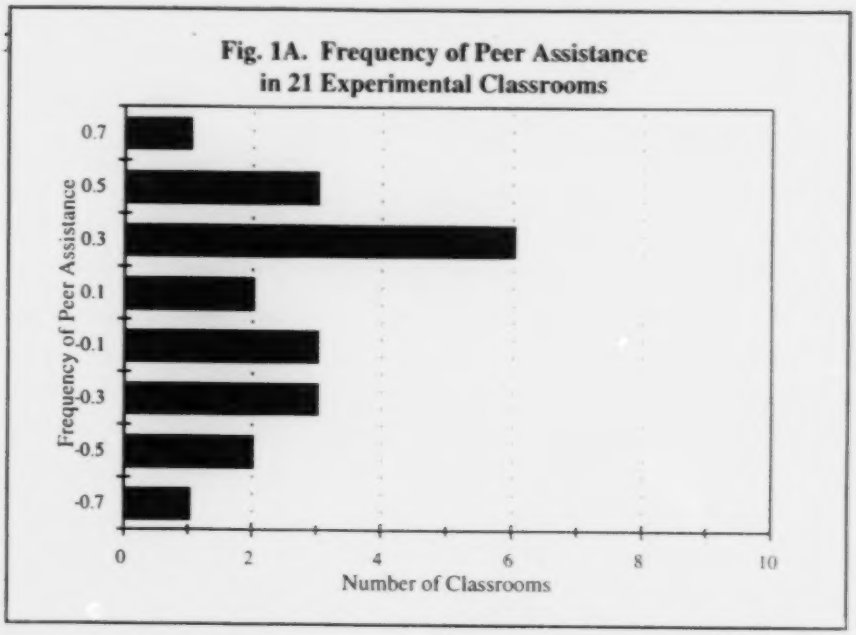
The multiple regression model summarized in Table 1 used school as a predictor of peer assistance in the classroom while controlling for the grade level of the classroom. This analysis indicated that peer assistance occurred .71 standard deviations more frequently in the typical experimental than in the typical control classroom. (The analysis also indicated that peer assistance is observed much more frequently in sixth grade classes than in eighth grade classes.) The histograms in Figures 1A and 1B depict the number of classrooms at each school that displayed each observed level of peer assistance. These figures provide a visual representation of the substantially higher frequency of peer assistance that was observed in most experimental classrooms than in most control classrooms.

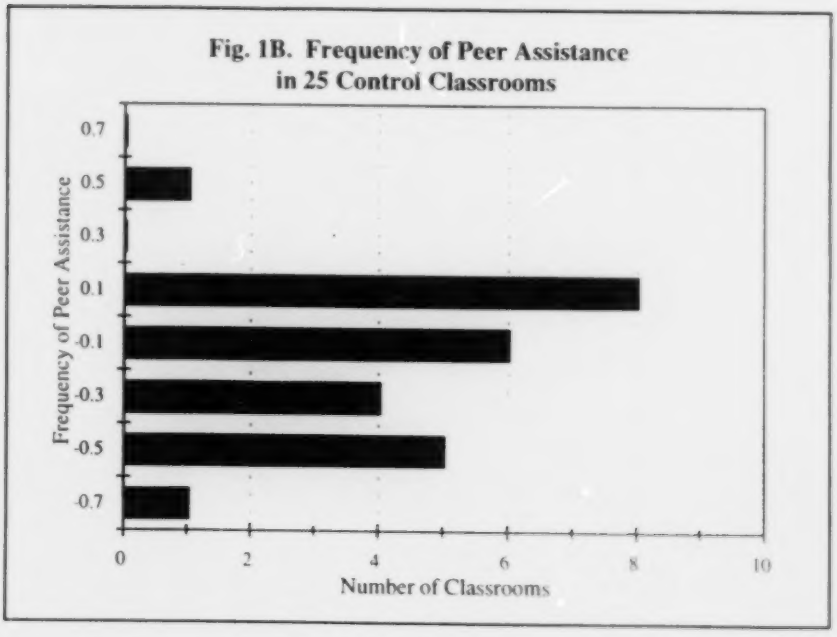
Table 1
Multiple Regression Model to Predict Frequency of Peer Assistance (N=46 classes)

Predictor	*b*	*se*	*p-value*	Effect Size
Intercept	-0.06	0.05	0.18	
School	0.22	0.09	0.02	0.71
Sixth Grade	0.40	0.11	0.001	1.30
Seventh Grade	0.18	0.11	0.12	0.58

Note: All predictors are grand-mean-centered. Total variance explained=0.30.

Frequency of Peer Assistance in Two Schools





19 BEST COPY AVAILABLE

School as a Predictor of Between-Classroom Differences in Students' Reading Comprehension

A matched control group, pretest-posttest design (Fitz-Gibbon & Morris, 1987) was used to evaluate school effects on students' end-of-year reading comprehension scale scores on the Stanford 9 after controlling for differences in pretest status (total reading scale scores on the CTBS from spring of the prior year). The conventional way of evaluating school effects in such a design is to conduct an ANCOVA at the student level, thus ignoring the fact that students are nested within classrooms. Fortunately, the development of hierarchical linear modeling (HLM) techniques makes it possible to estimate an ANCOVA-like model without erroneously assuming independent responses within classes (Bryk & Raudenbush, 1992).

We present the estimated parameters for two HLM models. Model 1 was specified as follows:

$$
\begin{aligned}
& Y_{ij} = \beta_{0j} + \beta_{1j}\,(\text{Prior Achievement}_{ij}) + r_{ij} \\
(1) \quad & \beta_{0j} = \gamma_{00} + \gamma_{01}\,(\text{School}_j) + \gamma_{02}\,(\text{Sixth grade}_j) + \gamma_{03}\,(\text{Seventh grade}_j) + u_{0j} \\
& \beta_{1j} = \gamma_{10} + u_{1j}
\end{aligned}
$$

Model 1 has an intercept, a measure of prior reading achievement, and a randomly varying error term at level 1, the student level. At level 2, the adjusted classroom mean level of achievement is modeled as a function of an intercept, school (1=Central East; 0=control site), grade (with eighth as the excluded reference category), and a randomly varying error term. The effect of prior achievement is modeled as a function of an intercept and a randomly varying error term. Thus, Model 1 provides us with an initial look at whether reading achievement in Central East Middle School's classrooms differed significantly from reading achievement in the control schools classrooms in Spring of 1996, controlling for students' prior achievement and controlling for grade levels.

Table 2 shows the results for this model. The coefficient for school (9.02) indicates that reading comprehension achievement was much higher at Central East Middle School than at the control site. Also note the positive and significant effect of prior achievement (0.40). We will not discuss the results of this initial model further because Model 2, described next, is a better-fitting, more appropriate model.

Table 2

Hierarchical Linear Model 1 for Reading Comprehension Achievement

Fixed Effects	Coefficient	*se*	*p-value*
Model for level-1 intercept, β_{0j}			
Intercept, γ_{00}	644.85	1.76	0.000
School, γ_{01}	9.02	3.23	0.008
Sixth Grade, γ_{02}	-20.95	4.18	0.000
Seventh Grade, γ_{03}	-6.35	3.79	0.101
Model for prior achievement slope, β_{1j}			
Intercept, γ_{10}	0.40	0.03	0.000

Random Effects	Variance Component	*df*	χ^2	*p-value*
Classroom mean, u_{0j}	90.25	40	143.74	0.000
Prior achievement slope, u_{1j}	0.02	43	96.47	0.000
Level-1 effect, r_{ij}	525.88			

The second model (2) adds just one thing to the first: school as a predictor of the effect of prior achievement on students' reading comprehension. Opponents of detracking and instructional programs like the Student Team Literature Program which are designed to provide a demanding curriculum to *all students* in *heterogeneous* RELA classrooms fear that students with the best prior achievement will suffer. If this fear is justified, the effect of prior achievement on students' reading comprehension should be weaker in experimental than in control classrooms. On the other hand, proponents of Student Team Literature argue that top achievers — because of their strong prior skills — will benefit even more than the average student from the challenging curriculum and state-of-the-art pedagogy that characterizes the Student Team Literature program. If this counter-argument is justified, the effect of prior achievement on students' reading comprehension should be stronger in experimental than in control classrooms.

$$Y_{ij} = \beta_{0j} + \beta_{1j}\,(\text{Prior Achievement}_{ij}) + r_{ij}$$

(2) $$\beta_{0j} = \gamma_{00} + \gamma_{01}\,(\text{School}_j) + \gamma_{02}\,(\text{Sixth grade}_j) + \gamma_{03}\,(\text{Seventh grade}_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\,(\text{School}_j) + u_{1j}$$

Table 3 summarizes the results of Model 2. The estimated impact of school on the reading comprehension of a "typical" student (one who is at the grand mean for grade level and prior achievement) is even stronger (11.7 points) in this model than in the previous one. In substantive terms, how large is this reading comprehension advantage we are seeing for the typical student at Central East? One way of gauging the size of this advantage to the typical student of being in a classroom at Central East rather than at the control site is to compare it to the size of the between-classroom standard deviation of the reading comprehension measure. This standard deviation is 22.78. Calculating an effect size by dividing the coefficient from Table 2 by this standard deviation reveals an effect size of 0.51 (11.70/22.78=0.51).

Another way to understand the substantive significance of Central East's reading comprehension advantage is to compare 11.7 to the reading comprehension advantage that a typical eighth grader has over a typical seventh grader in this sample. That difference, as shown in Table 3, is estimated to be 6.44 points. Thus, the advantage enjoyed by a Central East student over a student from the control school is almost twice the advantage an eighth grade student holds over a seventh grade student.

The estimate of γ_{11} in Table 3 indicates that the effect of prior achievement on students' reading comprehension is larger in the experimental classrooms than in the control classrooms. Specifically, the association between achievement on the 1995 test and the 1996 test was 30 percent stronger at Central East than at the control site (0.52/0.40=1.30). For top achievers — students whose prior achievement was one standard deviation above the mean on the CTBS in 1995 — students at Central East outperformed students at the comparison school by 17 points (substantially larger than the 11.7 point advantage for typical students). When this school effect on the prior achievement slope is considered together with the very strong school effect on the level-1 intercept (mean reading comprehension achievement in the classroom), the conclusion is that all students at Central East benefitted on average during the 1995-1996 school year, and that the top-achieving students benefitted especially.

Table 3
Hierarchical Linear Model 2 for Reading Comprehension Achievement

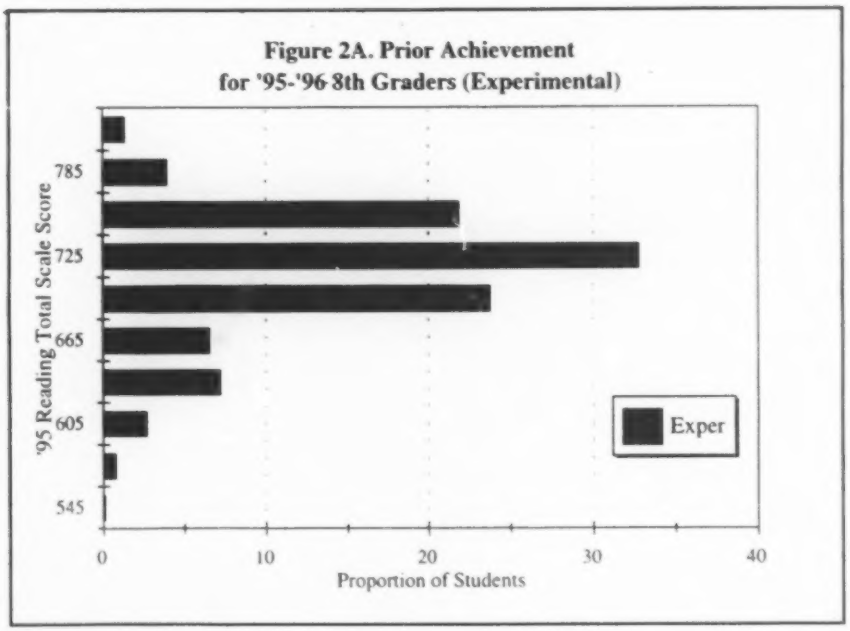
Fixed Effects	Coefficient	*se*	*p-value*
Model for level-1 intercept, β_{0j}			
Intercept, γ_{00}	644.83	1.77	0.000
School, γ_{01}	11.70	3.59	0.003
Sixth Grade, γ_{02}	-21.50	4.24	0.000
Seventh Grade, γ_{03}	-6.44	3.85	0.101
Model for prior achievement slope, β_{1j}			
Intercept, γ_{10}	0.40	0.03	0.000
School, γ_{11}	0.12	0.06	0.045

Random Effects	Variance Component	*df*	χ^2	*p-value*
Classroom mean, u_{0j}	92.03	40	141.63	0.000
Prior achievement slop, u_{1j}	0.02	42	82.52	0.000
Level-1 effect, r_{ij}	527.20			

The school effects revealed in our analyses can be seen in simple histograms that show the prior achievement distributions and the end-of-year achievement distributions at each school (holding grade level constant and including only students who were tested at both time points). For example, Figures 2A through 2D show the prior achievement distribution for eighth graders from Central East and the control site (on the 1995 CTBS test of reading achievement) and the end-of-year achievement distribution (on the 1996 Stanford 9 test of reading comprehension achievement).[3] In Figures 2A and 2C, data for 143 Central East

[3] Because the estimated school effect on reading comprehension achievement did not vary significantly by grade level, one obtains histograms regardless of which grade level is plotted. Histograms for sixth and seventh graders are available from the authors upon request.

Prior Achievement and End-of-Year Achievement for 8th Graders in Two Schools



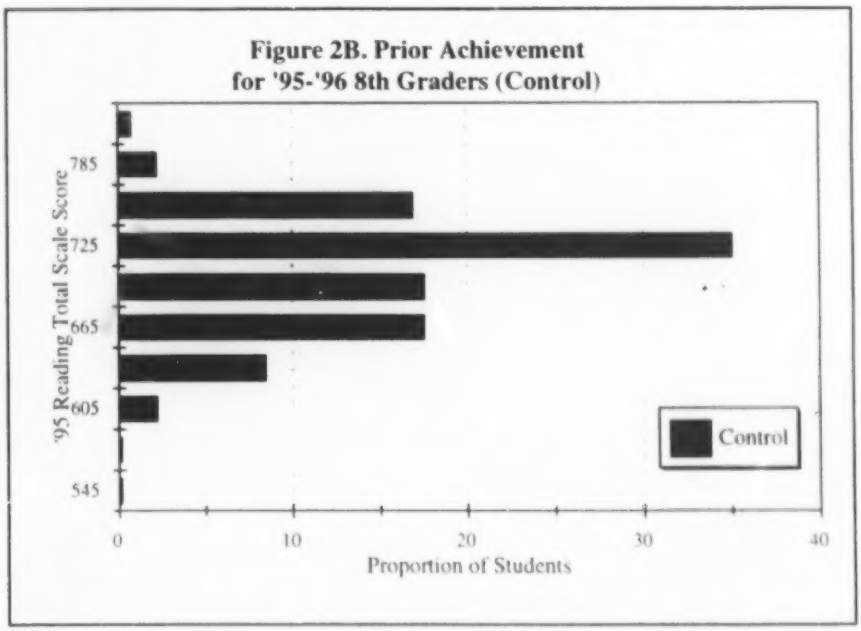


Figure 2C. End-of-Year Achievement for '95-'96 8th Graders (Experimental)

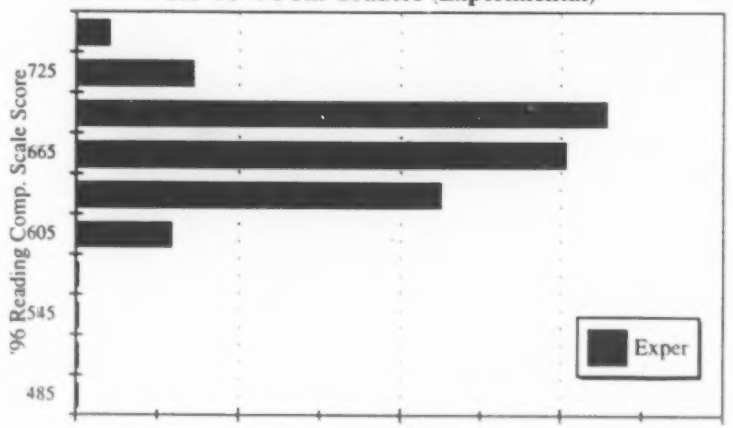
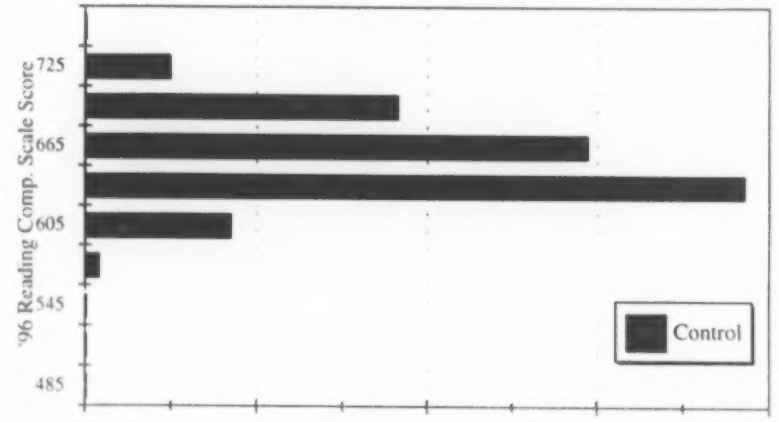


Figure 2D. End-of-Year Achievement for '95-'96 8th Graders (Control)



BEST COPY AVAILABLE

students are plotted at the two points in time; in Figures 2B and 2D, data for 156 students from the control site are plotted at the two points in time. What one sees are similar distributions at the two schools in 1995, but somewhat differing distributions in 1996. Specifically, in 1995, both schools had few students in the highest achievement range. The proportions of students in the lowest achievement ranges were similar with Central East, in fact, having a few students achieving at a level lower than that of any students from the control site. Finally, the modes for the two schools were found in the same thirty-point range.

By 1996, however, the distribution for Central East eighth graders tended toward the higher levels of achievement much more than the other school's distribution did. Some Central East students, but no students from the control school, are observed in the highest achievement range. Conversely, some students from the control site are observed in a low achievement range where no Central East students are observed. Finally, the modes for the two schools are clearly at different points. Thus, we take the estimates from Model 2 in Table 3 and this plotting of observed distributions as consistent and reinforcing evidence of the fact that there was a general achievement advantage at Central East and that top achievers (students with the highest prior achievement scores on the CTBS) benefitted especially.

School as a Predictor of the Productivity of Peer Assistance

Finally, Model 3 allows us to estimate the productivity of peer assistance in each school by adding the frequency of peer assistance and a frequency of peer assistance by school interaction term to Model 2:

$$Y_{ij} = \beta_{0j} + \beta_{1j}\,(\text{Prior Achievement}_{ij}) + r_{ij}$$

(3) $$\beta_{0j} = \gamma_{00} + \gamma_{01}\,(\text{School}_j) + \gamma_{02}\,(\text{Sixth grade}_j) + \gamma_{03}\,(\text{Seventh grade}_j) + \gamma_{04}\,(\text{Peer Assistance}_j) + \gamma_{05}\,(\text{Peer Assistance}_j * \text{School}_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\,(\text{School}_j) + u_{1j}$$

Table 4 displays the parameter estimates for Model 3. γ_{05} is an estimate of the differential productivity of peer assistance in experimental and control classrooms. Its magnitude (16.68) suggests that peer assistance is substantially more productive at Central East than at the control site.[4] Figure 3 illustrates these estimated effects of peer assistance on reading comprehension for the two schools. This figure suggests that peer assistance was

[4] Following usual standards (e.g., Pedhazur, 1982, p.440), a .10 level of significance was used in testing interactions.

productive at Central East Middle School and counterproductive at the control school. Based on the parameter estimates in Table 4, the predicted gain of a student in a class at the experimental school that moves up one standard deviation on the peer assistance scale is 1.9 points on the reading comprehension scale, the predicted decline for the same movement at the control school is 4.0 points. This raises a number of important issues about between-school differences in how peer assistance is organized which will be discussed shortly.

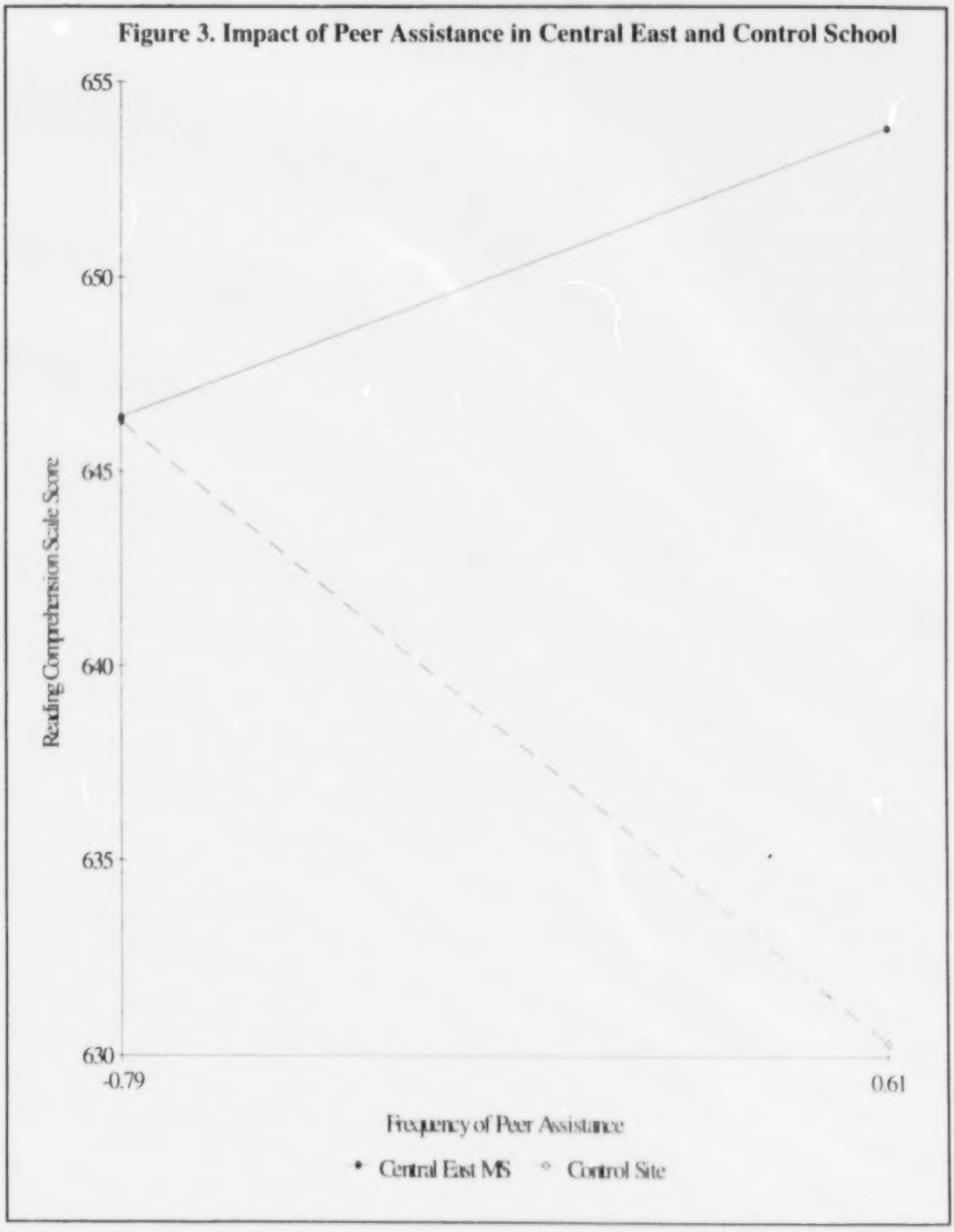
Table 4
Hierarchical Linear Model 3 for Reading Comprehension Achievement

Fixed Effects	Coefficient	*se*	*p-value*
Model for level-1 intercept, β_{0j}			
Intercept, γ_{00}	644.39	1.71	0.000
School, γ_{01}	13.25	3.80	0.002
Sixth Grade, γ_{02}	-19.77	4.81	0.000
Seventh Grade, γ_{03}	-4.10	3.95	0.306
Freq. of Peer Assistance, γ_{04}	-11.42	7.72	0.147
(Freq. of Peer Assist) * (School), γ_{05}	16.68	9.72	0.093
Model for prior achievement slope, β_{1j}			
Intercept, γ_{10}	0.40	0.03	0.000
School, γ_{11}	0.14	0.05	0.013

Random Effects	Variance Component	*df*	χ^2	*p-value*
Classroom mean, u_{0j}	85.87	39	133.44	0.000
Prior achievement slop, u_{1j}	0.01	43	73.68	0.003
Level-1 effect, r_{ij}	530.30			

BEST COPY AVAILABLE

HLM-Estimated Impact of Peer Assistance on Reading Comprehension in Experimental (Solid Line) and Control (Dashed Line) Classrooms



Discussion and Conclusions

After one year of implementation of the Student Team Literature program, a significant advantage in reading comprehension was observed at the first Talent Development Middle School relative to a closely matched comparison site. The improvements in reading comprehension at Central East were substantial. The observed effect size of .51 is quite large compared to the impact of other educational reforms and is of the magnitude needed to achieve serious academic gains (Mosteller, Light, & Sachs, 1996). The data also indicate that it was not just the "typical" student who improved. The histograms presented in figures 2A to 2D show that the reported mean gains appear, in part, to be the result of increasing the number of students in the upper reaches of the achievement distribution and decreasing the number of students in the lower reaches. Other analyses show that it was the high achievers who had the greatest absolute gains in reading achievement. In short, improvements in reading comprehension were widespread. The Student Team Literature instructional program in the context of the Talent Development Middle School Model appears, on average, to have benefitted not just low, middle, or top achievers but rather to have demonstrated a "common school effect" (Bryk, Lee, & Holland, 1993). Almost all of the students at Central East felt its positive impact.

The results of any quasi-experimental evaluation study need to be interpreted with caution. In interpreting the present study, one must remember that the results are for one year in forty-six classrooms. It is impossible to tell from this one evaluation study alone whether the dramatic advantage of being in a classroom that uses Student Team Literature will generalize to other classrooms in other schools or whether the classrooms at Central East Middle School will be able to sustain such positive results over multiple years. However, converging evidence from previous research documenting the effectiveness of related methods in other middle schools and elementary schools suggests that these results *are* generalizable (e.g., Stevens & Durkin, 1992; Stevens & Slavin, 1995).

Another reason for caution in interpreting these results is that the observed school effects on reading comprehension and on the effectiveness of peer assistance conceivably could be due to some systematic difference between the experimental and control classrooms other than the implementation of Student Team Literature's curriculum, instructional strategies, assessment practices, and professional development. Repeated visits to both the experimental and the control school by the Talent Development Middle School staff, however, give us several reasons to doubt the existence of a systematic difference between the schools on unmeasured confounding variables that would have an impact on reading

comprehension. Classroom visits at both schools indicate that each school has a fairly typical distribution of teaching talent. The number of gifted teachers is matched by an equal number who struggle. The classroom and hallway climate at both schools is generally good and in many respects very strong for inner-city middle schools. The amount of time typically allocated to RELA each day (90 minutes) is the same in both schools. Both schools emphasize the use of novels in RELA and we have observed a tremendous overlap in the novels selected for study at both schools. Both schools are considered desirable teaching assignments and thus are often on the receiving end of "voluntary transfers," teachers who give up seniority at another school to fill a vacancy at one of these two schools. In fact, the only two obvious differences we observed between the schools are that (a) Central East's classrooms are in a converted union hall and a converted factory, so that the resulting classroom space is often less optimal than that found in the control school (e.g., rooms are atypically small or of an atypical shape), and (b) Central East Middle was founded recently (in September 1992) and some of the school's original staff still teach at the school which may contribute to a sense of "ownership" of the school's students and programs. It is unlikely that either of these differences would be responsible for the large observed differences between the schools in students' gains in reading comprehension.

While the data we collected clearly show variation across classrooms at the experimental school in the level of Student Team Literature implementation, classroom visits also indicate that all the teachers were using our curriculum and assessment materials and at least some of the recommended instructional practices. Finally, although our numerous and at times surprise classroom visits by multiple staff members did witness adaptations of the Student Team Literature program at the classroom level, consistent or widespread use of alternative reading comprehension strategies or materials was not observed.

Although societal and economic constraints made it impossible for us to conduct the ideal evaluation study (a study in which researchers would randomly determine which students would attend Central East and which students would attend the control site), the evaluation design used here has many strengths. The reported gains in reading comprehension were found by comparing students who had the same levels of prior achievement and attended a demographically matched middle school in the same district. This design along with the use of HLM mitigates many of the factors which have traditionally confounded research on academic achievement. The reading comprehension gains also occurred after only one year of implementation. Thus the possibility that the reported results are somehow inflated by model misspecification (the absence of an important missing factor) needs to be balanced by the fact that the reported gains may have been deflated by the glitches of an implementation

year. It is likely that greater gains will be observed as teachers become more skilled and familiar with the materials and instructional strategies, students become more accustomed to peer assistance, and the full use of Student Team Literature spreads to the subset of teachers who were initially hesitant. Classroom visits during the 1996-97 school year, for instance, indicate that some of the low implementers in year one have subsequently become active users of Student Team Literature.

Clearly, continual and expanded evaluation of the Student Team Literature program in the context of the Talent Development Middle School model is needed. The results reported here need to be confirmed across multiple years in multiple schools before any final conclusions on its effectiveness can be drawn. The strength of these initial results, however, and the fact that they were obtained in a detracked inner-city middle school, lead us to believe that it is not too early to discuss some of the broader implications of the initial success of the Student Team Literature program at Central East Middle School.

The initial results of the Student Team Literature program appear to strongly support the small but growing literature on the academic benefits of peer-assisted learning (Fuchs et al., 1997). Our results, however, also clearly indicate that there may be better and worse ways to organize peer assistance. This is dramatically highlighted in the radically different estimated effects of peer assistance on reading comprehension in our experimental and control schools. At Central East an increase in the frequency of peer assistance is associated with an increase in reading achievement. At the control school, the opposite relationship was observed. There, more peer assistance was associated with lower achievement. Although we have not yet systematically investigated the nature of peer assistance at our control school, we suspect that the difference between our experimental and control school reflects the distinction between well-planned and structured peer-assisted learning (which increases both the number of students actively engaged with learning and the quality of their intellectual interactions) and relatively unorganized group or partner work. All available evidence indicates that it is not group or cooperative work *per se* which improves academic performance but what students do when they work with their peers. Peer-based activities which promote guided cooperative questioning and elaborative discussion, such as Student Team Literature, improve achievement, while activities which promote non-elaborative help or unguided questioning do not (King, 1994; Webb & Farivar, 1994). There is also ample evidence that peer learning which occurs within a context of group goals and individual accountability (as it does in Student Team Literature) is much more productive than peer learning which does not (Slavin, 1995). Finally, the success of peer assistance is linked to the time and effort spent preparing students for their roles and responsibilities (Williams, Harris,

& Hayakawa, 1995). It is an unfortunate and troubling trend that cooperative learning is increasingly being used as an umbrella term for both planned, structured, and facilitated peer assistance and relatively unorganized group work. The former can be a powerful multiplier of student achievement; the latter, as our results indicate, may be educationally unsound.

Our results also support the contention that it is possible to implement a demanding core curriculum in a detracked school and to create a situation in which students at all levels of the achievement distribution make significant and substantial gains in academic performance. At Central East, top achieving students did not suffer; in fact, they thrived. The existing literature suggests that their large achievement gains may have been the result of both access to rich and demanding curricular materials and the opportunities engendered by peer-assisted learning to explain and elaborate their thinking (King, 1994; Stodolsky, 1988).

In addition, the reported results for the Student Team Literature Program were obtained in an inner-city middle school. Knapp (1995) found that "teaching for meaning" was least likely to occur in classrooms that had class sizes greater than 27, high mobility, and concentrated poverty. The class size at Central East Middle is 33, and 85 percent of its students are eligible for free lunches. The strong gains in reading comprehension provide an initial validation of the Talent Development Middle School Model and its belief that a challenging instructional program, sufficiently supported and implemented schoolwide, can lead to large and widely distributed achievement gains in schools which serve students placed at risk.

Finally, the early results of the Student Team Literature program on middle school students' reading comprehension leave us hopeful that the challenge raised by the leaders of the major middle school reform initiatives can be met. It will not be easy. School reform is difficult and our own experience confirms that it requires constant attention at multiple levels (Useem, Christman, Gold, & Simon, 1997). The results, however, indicate that academically challenging instructional programs can be implemented and improve student performance across the board in the middle grades, if all students are provided with a demanding curriculum which promotes active learning and the accompanying instructional program is supported by structured peer assistance, sustained staff development, and assessments which provide useful feedback and facilitate student effort.

References

American Federation of Teachers (1992). U.S. education: The task before us. *American Educator, 16*(4), 19-30.

Balfanz, R. (1997, March). *"Mathematics for all" in two urban schools: A view from the trenches.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Beaton, A.E., Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Kelly, D.L., & Smith, T.A. (1996). *Mathematics achievement in the middle school years.* Chestnut Hill, MA: Boston College's TIMS International Study Center.

Bloom, B.S. (Ed.) (1984). *Taxonomy of educational objectives.* White Plains, NY: Longman.

Bruer, J.T. (1994). *Schools for thought: A science of learning in the classroom.* Cambridge, MA: MIT Press.

Bryk, A.S., & Raudenbush, S.W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Bryk, A.S., Lee, L.E., & Holland, P.B. (1993). *Catholic schools and the common good.* Cambridge, MA: Harvard University Press.

Cohen, D.K., Wilson, S., & Hill, H. (1997, March). *Teaching and learning mathematics in California.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Corbett, H.D., & Wilson B.L. (1997). *Urban students' perspectives on middle school: The sixth grade year in five Philadelphia middle schools.* Philadelphia, PA: Philadelphia Education Fund.

Elmore, R.F., Peterson, P.L., & McCarthey, S.J. (1996). *Restructuring in the classroom: Teaching, learning, and school organization.* San Francisco, CA: Jossey-Bass.

Fitz-Gibbon, C.T., & Morris, L.L. (1987). *How to design a program evaluation.* Newbury Park, CA: Sage.

Forgione, P. (1990). *Accountability and assessment: The Connecticut approach.* Paper presented at the Public Education Forum, Baltimore, MD.

Fuchs, D., Fuchs, L.S., Mathes, P.G., & Simmons, D.C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal, 34*, 174-206.

Glaser, R., & Silver, E. (1994). Assessment, testing, and instruction: Retrospect and prospect. *Review of Research in Education, 20,* 393-419.

Jones, L. (1996). *Listening comprehension: Secondary level.* Baltimore, MD: Johns Hopkins University, Center for Social Organization of Schools.

Jones, L. (1997). *Student Team Literature teacher's manual.* Baltimore, MD: Johns Hopkins University, Center for Social Organization of Schools.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31*(2),338-368.

Knapp, M.S. (1995). *Teaching for meaning in high-poverty classrooms.* New York, NY: Teachers College Press.

Lipsitz, J., Mizell, M.H., Jackson, A.W., & Austin, L.M. (1997, March). Speaking with one voice: A manifesto for middle grades reform. *Phi Delta Kappan,* 533-540.

Locke, E.A., & Latham, G.P. (1990). *A theory of goal setting and task performance.* Englewood Cliffs, NJ: Prentice Hall.

Mac Iver, D.J., &. Plank, S.B. (1998). "Improving urban schools: Developing the talents of students placed at risk." In J. Irvin (Ed.), *What current research says to the middle level practitioner* (pp.243-256). Columbus, OH: National Middle School Association.

Mac Iver, D.J. & Reuman, D.A. (1993/1994, Winter). Giving their best: Grading and recognition practices that motivate students to work hard. *American Educator*, 24-31.

Mac Iver, D.J., Reuman, D.A., & Main, S. (1995). Social structuring of the school. In M. R. Rosenzweig and L. W. Porter (Eds.), *Annual Review of Psychology: Vol. 46*, (pp. 375-400). Palo Alto, CA: Annual Reviews, Inc.

Madhere, S., & D.J. Mac Iver. (1996). *The Talent Development Middle School: Essential components.* Baltimore, MD & Washington, DC: Johns Hopkins University & Howard University, Center for Research on the Education of Students Placed at Risk.

Means, B., Chelemer, C., & Knapp, M.S. (Eds.) (1991). *Teaching advanced skills to at-risk students.* San Francisco, CA: Jossey-Bass Publishers.

Mevarech, Z.R., & Kramarski, B. (1997). IMPROVE: A multidimensional method for teaching mathematics in heterogeneous classrooms. *American Educational Research Journal, 34*(2), 365-394.

McLaughlin, M.W., & Oberman I. (Eds.) (1996). *Teacher learning: New policies, new practices*. New York, NY: Teachers College Press.

Mosteller, F.D., Light, R.J., & Sachs, J.A. (1996). Sustained inquiry in education: Lessons from skill grouping and class size. *Harvard Education Review, 66*(4), 797-842.

Natriello, G., McDill, E.L., & Pallas, A.M. (1990). *Schooling disadvantaged children: Racing against catastrophe*. New York, NY: Teachers College Press.

Nelson, B.S., & Hammerman, J.K. (1996). Reconceptualizing teaching: Moving toward the creation of intellectual communities of students, teachers, and teacher educators. In M.W. McLaughlin and I. Oberman (Eds.), *Teacher learning: New policies, new practices*. New York, NY: Teachers College Press.

Oakes, J., Gamoran, A., & Page, R. (1992) Curriculum differentiation: Opportunities, outcomes, and meanings. In P.W. Jackson (Ed.), *Handbook of research on Curriculum* (pp. 570-608). New York, NY: Macmillan.

Palincsar, A.S., & Klenk, L.J. (1991) Dialogues promoting reading comprehension. In B. Means, C. Chelemer, and M. Knapp (Eds.), *Teaching advanced skills to at-risk students*. San Francisco, CA: Jossey-Bass.

Pedhazur, E.J. (1982). *Multiple regression in behavioral research (2nd ed.)*. New York, NY: Holt, Rinehart, & Winston.

Porter, A. (1993). *Opportunity to learn* (Brief No. 7). Madison, WI: Center on Organization and Restructuring of Schools.

Schmidt, W.H., McKnight, C.C., & Raizen, S.A. (1996). *A splintered vision: An investigation of U.S. science and mathematics education*. Boston, MA: Kluwer.

Slavin, R.E. (1995). *Cooperative learning: Theory, research, and practice* (2nd ed.). Boston, MA: Allyn and Bacon.

Stevens, R.J., & Durkin, S. (1992). *Using Student Team Reading and Student Team Writing in middle schools*. Baltimore, MD: Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students.

Stevens, R.J., Hammonds, J., & Durkin, S. (1991). *Student Team Writing teacher's manual*. Baltimore, MD: Johns Hopkins University, Center for the Social Organization of Schools.

Stevens, R.J., & Shaw, A.M. (1991). *Listening comprehension*. Baltimore, MD: Johns Hopkins University, Center for Social Organization of Schools.

Stevens, R.J., & Slavin, R.E. (1995). The cooperative elementary school: Effects on students' achievement, attitudes, and social relations. *American Educational Research Journal, 32*(2), 321-351.

Stevenson, H.W., & Stigler J.W. (1992). *The learning gap*. New York, NY: Summit.

Stevenson, H.W., & Lee, S. (1995). The East Asian version of whole-class teaching. *Educational Policy, 9*(2), 152-168.

Stodolsky, S. (1988). *The subject matters*. Chicago, IL: University of Chicago.

U.S. Department of Education, National Center for Education Statistics (1996). *Pursuing excellence*. Washington, DC: Government Printing Office.

Useem, E.L., Christman J.B., Gold, E., & Simon, E. (1997). Reforming alone: Barriers to organizational learning in urban school change initiatives. *Journal of Education for Students Placed at Risk, 2*(1), 55-78.

Webb, N.M., & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal, 31*(2), 369-395.

Williams, D.R., Harris, J., & Hayakawa, C. (1995). Cooperative learning and conflict resolution: Perspectives of urban middle school adolescents. *Research in Middle Level Education, 18*(2), 23-47.

Appendix

Student Questionnaire Items in the Peer Assistance and Discussion Scale (a=.93)

Please tell us how often the following things happened in RELA or Reading class.

- Students took turns with partners asking questions, and answering the questions the partners asked.
- Students discussed a novel with a partner.
- Students worked in teams to master the vocabulary used in a novel.
- Students explained answers to their teammates and checked to make sure that all their teammates understood the material.

Note: To measure how often each learning activity occurred in each RELA classroom, we computed the mean response (in a z-score metric) of all the students in the class. First, student responses were coded in terms of school days per month (i.e., *never*=0 days per month, *once or twice a month*=1.5 days per month, *once or twice a week*=6 days per month, *most days*=12 days a month, *every day*=20 days per month). Z-scores were then calculated for each of the items. Then, the classroom mean z-score for each item was computed. Finally, a scale score for each class was computed by averaging classroom mean z-scores for the four items in the scale.

CRESPAR

Johns Hopkins University
Center for Social Organization of Schools
3505 North Charles Street
Baltimore MD 21218
410-516-8800 / 410-516-8890 fax

Howard University
2900 Van Ness Street, NW
Washington DC 20008
202-806-8484 / 202-806-8498 fax

BEST COPY AVAILABLE